DOCUMENT RESUME

ED 046 932 TE 002 179

AUTHOR Jurgens, Sister Jeanne Marie; Griffin, William J.
TITLE Relationships Between Overall Quality and Seven Language Features in Compositions Written in Grades Seven, Nine, and Eleven.
INSTITUTION George Peabody Coll. for Teachers, Nashville, Tenn. Inst. on School Learning and Individual Differences.
REPORT NO PP-70-01
PUB DATE 70
NOTE 23p.

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS Composition (Literary), *Composition Skills (Literary), Discourse Analysis, Evaluation Criteria, Evaluation Methods, Language Patterns, *Language Research, *Linguistic Patterns, Phrase Structure, Rating Scales, Secondary Education, Transformation Generative Grammar, Transformations (Language), *Writing Skills, *Written Language

ABSTRACT

Seven objective measures of language production--number of words, T-units, subordinate clauses, clauses of all types, words per clause, words per T-unit, and clauses per T-unit--were applied to 269 compositions written by seventh, ninth, and 11th graders to test assumptions about the relationships between quality of writing and students' age and maturity levels. Quality ratings of high (H), middle (M), and low (L) were given to each paper, using the evaluation procedures of the STEP test. The following relationships were tested: for total words, total T-units, words per T-unit, number of subordinate clauses, and total clauses, it was hypothesized that H=M, M>L, and H>L; for words per clause, H=M=L; and for clauses per T-unit, H<M, M>L, and H>L. Results indicated that (1) distinctions between quality of papers were not signaled by differences in ratios of words to T-units, words to clauses, or clauses to T-units, (2) quality was distinguished by significant increases from low to middle to high in total words, total T-units, and total clauses, and (3) on the whole, low papers made use of significantly fewer subordinate clauses than did either middle or high. (Included are tables indicating the statistical measures used and a review of normative studies which previously examined developmental trends in composition quality. (JM)

RELATIONSHIPS BETWEEN OVERALL QUALITY AND SEVEN LANGUAGE FEATURES IN COMPOSITIONS WRITTEN IN GRADES SEVEN, NINE, AND ELEVEN

Sister Jeanne Marie Jurgens and
William J. Griffin

Professional Paper No. 70-01

Institute on School Learning and Individual Differences
John F. Kennedy Center for Research on Education and Human Development
George Peabody College for Teachers Nashville, Tennessee 37203

Relationships Between Overall Quality and Seven Language Features in Compositions Written in Grades Seven, Nine, and Eleven*

By

Sister Jeanne Marie Jurgens and William J. Griffin

Various types of normative studies[1] have quite consistently identified developmental trends in the language production of children at successive age-grade levels. In what ways may measurable differences that generally distinguish the language performance of older children from that of younger be related to distinctions between better and poorer writing elicited under uniform (or at least very similar) stimulus conditions from children in a particular age range? This is the question on which the investigation reported here was intended to throw some light.

Even cursory consideration will suggest the irrationality of anticipating simple, consistent correlations of higher quality ratings with higher incidence rates of all the language features that develop progressively in the process of maturation. Leaving entirely out of account the probability that judges of composition quality do not operate uniformly in arriving at their evaluations, it is obvious that relative success in writing depends on numerous factors, some of which are not dealt with in the normative studies that have

---

*This study was supported in part by funds made available to the Institute on School Learning at George Peabody College for Teachers by the Carnegie Corporation of New York. The authors, however, are solely responsible for statements made and views expressed in the report.

1. Earlier studies are summarized by McCarthy (1954); some of the important recent studies are those of Harrell (1957), Strickland (1962); Loban (1963, 1966); Riling (1965); Hunt (1965); and O'Donnell, Griffin, and Norris (1967).

been referred to. Common observation shows that some valuable skills in writing may be offset by various compositional defects, and some weaknesses may be compensated for by other virtues.

The situation is dramatized in the report by Mellon (1969) of an admirably planned experiment in the seventh grade. Students in experimental classes were led to produce compositions displaying (on every one of 12 counts) a significantly higher degree of "maturity in sentence structure" than was to be found in comparable writing done in control classes. Yet, when representative samples of papers from the two groups (typed with corrections in spelling and punctuation and without author identification) were evaluated by well trained judges who were junior high school teachers uninformed about the experiment, the ratings given students in the control group were significantly higher. Having satisfied himself that the raters' judgments were both consistent and valid, and having considered that all students involved had done the same amount of composition writing during the year, Mellon (p. 70) concluded that "the differences in quality of writing resulted from the effect of teacher variable, that is, the fact that the highly experienced control teacher was a better composition instructor...."

The Mellon report is cautionary in several respects, but it does not directly answer the question to which this investigation is addressed--nor was it intended to do so. Only a few studies of the subject, indeed, have been reported.

Hillocks (1964) found that among the various objective syntactic measures he applied to ninth grade themes,[2] mean sentence length best discriminated

2. Other measures applied by Hillocks were (1) computation of sentences longer than 20, 30, and 40 words; (2) ratio of subordinate clauses to total clauses; (3) ratio of number of verbals to total sentences; and (5) relative number of sentences introduced by adverbial clauses.

the A, B, and C quality groupings to which they had been assigned by teachers. Griffin (1967) computed correlations of quality ratings and the single factor of T-unit[3] length in two groups of 177 randomly selected expository papers written in response to the same assignment in, respectively, a junior high school (Grades 7, 8, 9) and a senior high school (Grades 10, 11, 12). In neither group did he find a statistically significant difference in mean T-unit length of papers rated below average and those rated average and better, but the age range of the writers in each of his two groups casts doubt on the usefulness of his observation. Potter (1967) analyzed numerous syntactic characteristics of the 20 best and 20 poorest among 100 tenth grade expository compositions and was led to a number of conclusions by inspection of tabulated data. Since he did not subject his data to tests of statistical significance, and since most of his computations are reported as simple frequency counts (not reduced to ratios), interpretations at several points may be doubtful. Implications appear clear enough, however, in the findings that subordination indexes for the two groups of papers were exactly the same, that each group had the same percentage of T-units with prenominal openers, and that almost precisely the same number of adjectival clauses occurred in each. On the other hand, notable differences were observed in several features. The "good" papers made use of longer sentences and (in exactly the same degree) longer T-units; they contained twice as many passive constructions; they employed more prepositional phrases, and they showed more modifications of objects of prepositions; they

3. "The T-unit" is a convenient term invented by Hunt (1965) to label what is commonly called a simple or complex sentence. (Compound or compound-complex sentences are composed of two or more T-units).

used more "thought linking" transitional expressions as sentence openers; and they employed more verbals in each of four different functions. The "poor" papers contained more T-units, more clausal patterns of the forms S-V-O and S-V-Adv., more conditional clauses used as sentence openers, and somewhat more nominal and adverbial clauses. The investigator also noted that the poor writers were less proficient than the good writers in the use of coordinating words and devices.

In evaluation procedures as well as in selection of linguistic features for study, investigations that most closely resemble the one reported here are those of Veal (1969) and Biesbrock and Veal (1969). Veal (1969) collected eight papers each from subjects in Grades 2, 4, and 6. Papers from 27 randomly selected children in each of the three grades were evaluated on a seven-point scale, following the Educational Testing Service routine for rating STEP essays.[4] The quality ratings of all eight papers produced by a particular child were averaged, and on the basis of these averages the children in each grade were divided into equal groups designated as "high," "middle," and "low." Papers by children in each of the nine subgroups were then studied to acquire data on total word production, total T-units, word-length of T-units, the number of subordinate clauses used, total clauses produced, the ratio of all clauses to the number of T-units, and the frequency of coordinators used to join T-units. An additional analysis of subordinate clauses distinguished between those that were nominal, adverbial, and adjectival. Except for the count of words per clause and the count of T-unit coordinators, all the objective measures

4. See Handbook for Essay Tests, Level 1 (Princeton and Los Angeles: Educational Testing Service, 1957), pp. 5, 19-20.

applied marked significantly (at the .05 level or better) at least one quality group distinction in each grade; with the added exception of words per T-units in Grade 4, they all significantly marked at least two quality distinctions in each grade. But two important questions should be raised. First, did not the investigator confuse issues by summing quality ratings for each child? The result of this summing was that average writing behavior of individuals was described rather than linguistic characteristics of papers differentially rated for quality. Second, was the investigator well advised in making simple equal-group divisions of subjects in each grade? There is no reason to suppose that such a division will give the same ranges of ratings in the subgroups in various grades; and even if it does, the qualitative difference between poorer papers in one subgroup and better papers in the next lower subgroup is likely to be marginal indeed.

A different kind of question needs to be asked regarding the study reported by Biesbrock and Veal (1969). It clubbed together papers written by second and third graders, and hence, like the Griffin (1967) investigation, it did not focus on writing behavior at a single stage of maturation. From approximately 1,000 papers written in the two grades and evaluated according to the Educational Testing Service system, 60 were selected which had been given identical quality ratings by each of four judges. Five papers represented the top level and five papers the bottom level on the seven-point scale; each of the other levels was represented by ten papers. It was found that, while the mean T-unit length of the best papers was almost twice that of the poorest, the overall correlation between quality ratings and T-unit length was only .48. Higher correlation

coefficients were found between quality ratings and the number of T-units produced (.75), the ratio of total clauses to T-units (.62), the number of subordinate clauses produced (.71), and the number of all clauses (.80). On all of these measures, higher production rates were associated with higher evaluations of quality.

Design of the Present Study

In the study reported here, the plan was to apply seven objective measures of language production to seventh, ninth, and eleventh grade compositions that, in each of the grades, were judged to fall into distinctly differentiated quality groups, designated as "high," "middle," and "low." Four of the measures required simple group means of total production of words, T-units, clauses of all types, and subordinate clauses. The three remaining measures involved, for each group of papers in each grade, the computation of ratios: words per clause, words per T-unit, and clauses per T-unit. Analyses of variance and multiple comparisons of differences in quality group means were performed to test, at the .05 level of confidence, a set of relevant hypotheses.

Hypotheses to be tested reflected an assumption that most of the measures applied would not significantly discriminate High papers from Middle papers but would clearly distinguish Low papers from those in the other two groups. The assumption was that at a particular age level the poorest writing can be expected to show a characteristically lower degree of linguistic maturity, but that most of the important distinctions between good and better writing will be found in other aspects of compositional skill. It was, however, hypothesized that in words per clause no significant difference would be found in any of the quality groups. The reasoning was

that, while the more highly skilled writers would lengthen some of their clauses through use of subclausal modifications and coordinations, they would also be more likely to embed clause within clause, a practice that normally has the effect of reducing the mean length of total clauses employed. It was also anticipated that, though both High and Middle papers would show a higher ratio of clauses per T-unit than would be found in Low papers, the ratio would be significantly lower in the High than in the Middle group, because the best writers would more characteristically demonstrate syntactic control through employment of deletion transformations. Finally, it was anticipated that differential relations between quality groups and the objectively measured language features would be uniform in the three grades to be studied.

The relations hypothesized can be summarized as follows:

1. For total words, total T-units, words per T-unit, number of subordinate clauses produced, and total clauses

   H = M
   M > L
   H > L

2. For words per clause

   H = M = L

3. For clauses per T-unit

   H < M
   M > L
   H > L

Procedures

Compositions dealt with in this study were drawn from a pool of writing samples collected for a more comprehensive investigation. All students present

in the Coffee County, Tennessee, junior and senior high schools on a particular day in January, 1968, were asked to write papers entitled, "A Fictional Character I'd Like to Know in Real Life." Directions for the assignment were presented in mimeographed form, and teachers were instructed to give no assistance and to make no corrections. The writing session was limited to 30 minutes. Under identical limitations and following closely comparable directions, the students in the same schools four months later wrote another composition on a similar topic: "The Historical Character I Would Most Like to Know as a Friend." The original intent was to compare the work of equal numbers of boys and girls in the classes of each teacher in each grade at two writing times. From the January papers, therefore, compositions written by nine boys and nine girls in classes of each teacher at each grade level were randomly selected for evaluation and objective analysis. Papers written by most of the same students in May were later also evaluated and analyzed, but dropouts and absences somewhat reduced the number of second writing samples to be studied. In the original investigation, groups were once more equalized by random elimination of some of the January papers; but for the present study, which compares papers and not writers, it was conceived to be appropriate to work with all the seventh, ninth, and eleventh grade papers that had been evaluated and otherwise processed. For both the seventh and ninth grades these amounted to 36 January papers and 31 May papers (total, 67 each); for the eleventh grade, there were 72 January papers and 63 May papers (total, 135).

Quality ratings were assigned the papers by graduate students in English trained to use the procedures developed by the Educational Testing Service for evaluation of STEP essay tests. (See fn. 4 above.) In this

system, raters are instructed to consider the composition as a whole, but to keep in mind that "quality of content" should count about 50%, "style" about 30%, and "conventions" about 20%. A special feature of the ETS rater training is agreement in the selection of models that represent various degrees of global quality. One modification of the ETS system was made: the adoption of an eight-point (rather than a seven-point) scale, in which a "0" value was assigned papers judged to be thoroughly incoherent or totally inappropriate. Readers worked with typescripts of compositions that were identified only by code numbers, and were thus protected from halo effects of handwriting and irrelevant associations with names and sex identifications. The typescripts, however, meticulously preserved the original spelling, punctuation, capitalization and paragraph identations.

Each paper was evaluated by two readers working independently; the rating finally assigned a paper was the mean of the two scores. Papers from Grades 7, 8, and 9 and those from Grades 10, 11, and 12 were evaluated by different pairs of judges at each of the two writing times. Both in January and in May, models to exemplify quality levels 2, 4, and 6 were drawn separately from junior high and senior high papers that had been randomly excluded from the samples to be evaluated. Inter-rater reliabilities in evaluations of the total sample of papers studied were calculated by applying the Pearson Product Moment formula with the Spearman-Brown correction. The coefficients of correlation derived were as follows:

Junior High ratings in January	.92
Junior High ratings in May	.89
Senior High ratings in January	.93
Senior High ratings in May	.83

For the kinds of comparisons planned in the present study, it appeared desirable to work with distinctly differentiated quality groups of papers rather than with a finely graduated continuum. It seemed reasonable also to eliminate from consideration one seventh grade paper to which both readers had assigned a value of "0." "Low" papers, then were defined as those whose ratings were 0.5, 1.0, or 1.5; "Middle" papers were identified as those with ratings of 2.5, 3.0, or 3.5; "High" papers included all those with ratings of 4.5 or above. Thus, each quality group was separated from the one next to it in the scale by a full rating point. It will be noted that the possible range of ratings of papers designated "High" was wider than that of the other two groups; but the facts are that among the papers from Grades 7, 9, and 11, none was given a 7 rating, only one (from Grade 9) was rated 6.5, and only two (one from Grade 7 and another from Grade 9) were rated 6. Elimination of the "0" paper and those within the intervals separating the defined groups left for study 50 compositions from the seventh grade, 51 from the ninth, and 86 from the eleventh. The distribution of papers and the mean quality rating for each group in each of the three grades are presented in Table 1.

Linguistic analyses of the papers were performed on accurate typescripts by an analyst who had no knowledge of the quality ratings that had been assigned them. The analyst first marked for exclusion from computations all garbles (false starts, abnormal redundancies, and word tangles). In word counts, contractions were considered as two words, and compound nouns were given the count indicated by the number of morphemic bases in them; otherwise, individual words were taken to be identified by ordinary dictionary entries. The texts were segmented into T-units, following the

Table 1

Number of Papers and Average Quality Ratings at
High, Middle, and Low Levels In
Grades 7, 9, and 11

Grade	Level	Average Quality Rating	Number of Papers
7	High	5.12	8
9	High	6.00	3
11	High	4.78	9
7	Middle	2.96	28
9	Middle	2.91	17
11	Middle	2.99	43
7	Low	1.18	14
9	Low	1.16	31
11	Low	1.22	34

procedure established by Hunt (1965). (See fn. 3 above.) A clause was understood to be a syntactic structure composed of a grammatical subject and a predicate, either of which might involve coordination. The total number of clauses was computed by adding the number of subordinate clauses to the number of T-units. Each subordinate clause was given the count of 1, whether it operated on the first, second, third or lower level of subordination. (A first level subordination, which directly modifies an independent clause or some element in it, may itself be modified by a second level subordination, etc.) The ratios computed in this study were figured by dividing the total number of words by the total number of T-units, the total number of words by the total number of clauses, and the total number of clauses by the total number of T-units.

Findings

For each of the three grades, quality group means on seven linguistic measures are reported in Table 2. Inspection will show that on the measures of total words, total T-units, and words per clause there was, without exception in any grade, a clear increment in means for the Middle group as compared with the Low and for the High group as compared with the Middle. In the other categories, however, it will reveal that the investigators were wrong in assuming uniform performance patterns would be found in all grades. While relative differences in mean scores of quality groups in Grade 7 are generally paralleled by those in Grade 11, the Grade 9 groups reflect quite different interrelationships at several points.

The tests of statistical significance of difference in means are reported in Tables 3, 4, and 5. Inspected in conjunction with data shown in Table 2, they will indicate which of the investigators' hypotheses were or were not confirmed at the .05 level.

In all grades, there was confirmation of the hypotheses that Low papers would be shorter, would employ fewer T-units, and would contain fewer clauses than those in the other groups; but only in Grade 9 was there confirmation of the related hypotheses that on these counts no significant differences would be found between Middle and High papers. In Grades 7 and 11, High groups had significantly greater means in these categories than did Middle groups.

Table 2

Means on Seven Linguistic Measures for Each

Grade and Quality Level

Grade	Level	Total Words	Total T-units	Words per T-unit	Subordinate Clauses	Total Clauses	Words per Clause	Clauses per T-unit
7	H	173.50	13.62	13.25	8.00	21.62	8.10	1.66
	M	123.86	10.18	13.23	6.04	16.21	7.90	1.67
	L	80.29	7.50	11.34	4.14	11.64	7.21	1.53
9	H	147.00	11.67	12.93	5.00	16.67	8.75	1.47
	M	132.41	10.24	14.40	7.24	17.47	8.00	1.81
	L	73.77	5.77	13.13	4.65	10.58	7.12	1.87
11	H	271.67	18.89	16.47	13.89	32.78	9.35	1.76
	M	173.00	11.70	15.67	9.30	20.97	8.42	1.86
	L	86.85	6.21	14.72	4.65	10.85	8.24	1.80

Table 3

Multiple Comparison of Differences Between Means of Quality Groups in Grade 7

Variable	Quality Groups Compared		Mean Squares	F	P(F)
Total Words	H-M		15334.21	13.81	.0008
	M-L		17719.51	15.96	.0002
	H-L		44235.09	39.83	.0000
		Error	1110.52		
Total T-units	H-M		73.93	4.23	.0427
	M-L		66.93	3.83	.0266
	H-L		190.98	10.93	.0010
		Error	17.48		
Words per T-unit	H-M		0.00	0.00	.9848
	M-L		33.02	2.93	.0448
	H-L		18.41	1.63	.1023
		Error	11.26		
Subordinate Clauses	H-M		24.02	2.50	.1163
	M-L		33.44	3.49	.0323
	H-L		75.77	7.90	.0035
		Error	9.59		
Total Clauses	H-M		182.17	5.96	.0174
	M-L		195.09	6.39	.0071
	H-L		507.36	16.61	.0001
		Error	30.55		
Words per Clause	H-M		0.24	0.14	.7112
	M-L		4.52	2.59	.1101
	H-L		4.06	2.33	.1294
		Error	1.74		
Clauses per T-unit	H-M		0.00	0.01	.4665
	M-L		0.19	1.21	.1884
	H-L		0.08	0.54	.2645
		Error	0.16		

Table 4

Multiple Comparison of Differences Between Means of

Quality Groups in Grade 9

Variable	Quality Groups Compared		Mean Squares	F	P(F)
Total Words	H-M		542.73	0.37	.5544
	M-L		37749.64	25.52	.0000
	H-L		14666.77	9.92	.0015
		Error	1479.07		
Total T-units	H-M		5.22	0.30	.5925
	M-L		203.00	11.69	.0008
	H-L		89.83	5.17	.0129
		Error	17.37		
Words per T-unit	H-M		5.51	0.36	.5576
	M-L		17.62	1.15	.1438
	H-L		0.11	0.01	.4647
		Error	15.26		
Subordinate Clauses	H-M		12.73	1.60	.2095
	M-L		73.13	9.19	.0020
	H-L		0.32	0.04	.4172
		Error	7.96		
Total Clauses	H-M		1.64	0.05	.8179
	M-L		521.20	15.95	.0002
	H-L		101.31	3.10	.0405
		Error	32.68		
Words per Clause	H-M		1.43	0.65	.5727
	M-L		8.57	3.91	.0506
	H-L		7.31	3.34	.0705
		Error	2.19		
Clauses per T-unit	H-M		0.28	1.36	.1241
	M-L		0.04	0.23	.3191
	H-L		0.43	2.09	.0754
		Error	0.21		

Table 5

Multiple Comparison of Differences Between Means of

Quality Groups in Grade 11

Variable	Quality Groups Compared		Mean Squares	F	P(F)
Total Words	H-M		72416.93	24.99	.0000
	M-L		140986.94	48.66	.0000
	H-L		243065.15	83.89	.0000
		Error	2897.58		
Total T-units	H-M		384.84	14.46	.0005
	M-L		572.68	21.51	.0000
	H-L		1144.71	43.00	.0000
		Error	26.62		
Words per T-units	H-M		7.86	0.41	.5305
	M-L		9.84	0.51	.2587
	H-L		21.74	1.13	.1448
		Error	19.15		
Subordinate Clauses	H-M		156.52	8.40	.0049
	M-L		411.43	22.09	.0000
	H-L		607.70	32.63	.0000
		Error	18.62		
Total Clauses	H-M		1036.44	15.03	.0004
	M-L		1946.08	28.21	.0000
	H-L		3420.83	49.60	.0000
		Error	68.97		
Words per Clause	H-M		6.08	2.68	.1013
	M-L		0.79	0.35	.5636
	H-L		8.73	3.85	.0500
		Error	2.27		
Clauses per T-unit	H-M		0.07	0.40	.2669
	M-L		0.07	0.40	.2677
	H-L		0.01	0.05	.4053
		Error	0.19		

The hypothesis that High and Middle papers would not be significantly distinguished on the score of words per T-unit proved to be justified in each grade. But contrary to expectations, in no grade was the ratio of words to T-units significantly higher for High than for Low papers, and only in Grade 7 was the higher ratio for Middle papers as compared with Low papers significant. On the other hand, the hypothesis that there would be no significant group differences in words per clause was unconfirmed only in the comparison of High and Low papers in the eleventh grade, where High papers employed the longer clauses. (It should perhaps be noted, however, that the hypothesis was confirmed very marginally in the comparison of Middle and Low papers in Grade 9, where clauses in Middle papers were the longer ones.)

The expectations about production of subordinate clauses were entirely accurate in only one respect. In all grades, Middle papers contained a significantly larger number of such clauses than did Low papers. It had been anticipated that the same relation would hold between High and Low papers, but in Grade 9 the greater number of subordinate clauses in High papers was not significant. It had also been hypothesized that Middle and High papers would not be significantly differentiated by the numbers of subordinate clauses found in them, but this proved to be the fact only in Grades 7 and 9; in Grade 11, High papers had a significantly higher number of such clauses.

None of the hypotheses relating to clauses per T-unit was confirmed. It had been expected that Low papers would show a significantly lower ratio than would either High or Middle papers, but that the ratio in High papers would be significantly lower than in Middle papers. In point of fact, no

significant difference in this category was found in any of the comparisons.

A summary of the hypotheses tested and the outcomes of the tests is offered in Table 6.

Table 6

Summary of the Hypotheses

Confirmed (C) and Not Confirmed (NC)

Variable	Hypothesis	Grade 7		Grade 9		Grade 11	
		C	NC	C	NC	C	NC
Total Words	H = M		X	X			X
	M > L	X		X		X	
	H > L	X		X		X	
Total T-units	H = M		X	X			X
	M > L	X		X		X	
	H > L	X		X		X	
Words per T-unit	H = M	X		X		X	
	M > L	X			X		X
	H > L		X		X		X
Subordinate Clauses	H = M	X		X			X
	M > L	X		X		X	
	H > L	X			X	X	
Total Clauses	H = M		X	X			X
	M > L	X		X		X	
	H > L	X		X		X	
Words per Clause	H = M	X		X		X	
	M = L	X		X		X	
	H = L	X		X			X
Clauses per T-unit	H < M		X		X		X
	M > L		X		X		X
	H > L		X		X		X

Discussion

In this study of seven language features in compositions of High, Middle, and Low quality written in Grades 7, 9, and 11, perfect uniformity was found only in the unanticipated lack of significant differences in the ratios of clauses to T-units among the subgroups in every grade. It is also true, however, that in fourteen of the twenty-one comparisons of pairs of subgroups, the same relations were found in all grades. In performance relating to five of the language features analyzed, uniformity in the three grades was broken at only one of nine points of possible rupture, while in production of subordinate clauses there were two irregularities.

If we leave out of account all single exceptions, we may formulate a simple pair of generalizations that typify some aspects of language behavior reflected in the papers analyzed.

1. Distinctions between quality subgroups of papers were <u>not</u> signalized by differences in ratios of words to T-units, words to clauses, or clauses to T-units.
2. Quality subgroups <u>were</u> distinguished by significant increases from Low to Middle to High in total words, total T-units, and total clauses produced.

If we choose to ignore two exceptions, we may add the generalization that Low papers made use of significantly fewer subordinate clauses than did either Middle or High papers, which were themselves not distinguished by difference in the frequency of such clauses.

It is not certain, of course, that the population reported on here is reliably representative of writers in junior and senior high schools. Even if it is, there is a possibility that interrelationships among quality-rated

groups of compositions may be partly dependent on topics and rhetorical types of the papers written.

Finally, it should also probably be observed that when language features of ranked quality groups of compositions are compared, the patterns of relationships may differ with differing degrees of separation between the groups. In the present study it may be noted that three of the seven irregularities identified were found in comparisons of High and Middle papers in Grade 9. These irregularities might be dismissed on the ground that the number of High ninth grade papers was very small. But it is also true (as a glance at Table 1 will show) that the few Grade 9 papers rated High were considered by the evaluators to be considerably better than the High papers in Grades 7 and 11. Since relations between High and Middle ninth grade papers were consistently those hypothesized by the investigators (except that the lowered ratio of clauses to T-units in High papers did not quite reach statistical significance), it is tempting to assume that if the spread between High and Middle papers in Grades 7 and 11 had resembled that in ninth grade papers, the seventh and eleventh grade patterns of interrelationships would have paralleled those found in the ninth grade. Such an outcome would have been gratifying.

References

Biesbrock, Edieann F., and L. Ramon Veal. "A Study Comparing Global Quality and Syntactic Maturity in the Written Composition of Second and Third Grade Students." Unpublished paper read at the American Educational Research Association meeting, February, 1969.

Griffin, William J. "Development of Syntactic Control in Children's Writing." Unpublished paper read at the Conference on College Composition and Communication, April, 1967.

Harrell, Lester E., Jr. "A Comparison of Oral and Written Language in School-age Children." Monographs of the Society for Research in Child Development, 22, No. 3, 1957.

Hillocks, George. "An Analysis of Syntactic Patterns in Ninth Grade Themes." Journal of Educational Research, 57, 1964, pp. 417-419.

Hunt, Kellog W. Grammatical Structures Written at Three Grade Levels. National Council of Teachers of English Research Report No. 3. Champaign, Ill.: National Council of Teachers of English, 1965.

Loban, Walter D. Language Ability: Grades Seven, Eight, and Nine. Cooperative Research Monograph No. 18. Washington: U.S. Department of Health, Education, and Welfare, 1966.

Loban, Walter D. The Language of Elementary School Children. National Council of Teachers of English Research Report No. 1. Champaign, Ill.: National Council of Teachers of English, 1963.

McCarthy, Dorothea A. "Language Development in Children." In Manual of Child Psychology, second ed., edited by Leonard Carmichael. New York: John Wiley and Sons, 1954, pp. 492-630.

Mellon, John C. Transformational Sentence-Combining: A Method for Enhancing the Development of Syntactic Fluency in English Composition. National Council of Teachers of English Research Report No. 10. Champaign, Ill.: National Council of Teachers of English, 1969.

O'Donnell, Roy C., William J. Griffin, and Raymond C. Norris. Syntax of Kindergarten and Elementary School Children: A Transformational Analysis. National Council of Teachers of English Research Report No. 8. Champaign, Ill.: National Council of Teachers of English, 1967.

Potter, Robert R. "Sentence Structure and Prose Quality." Research in the Teaching of English, 1, 1967, pp. 17-28.

Riling, Mildred E. Oral and Written Language of Children in Grades 4 and 6 Compared with the Language of Their Textbooks. Report to the U.S. Office of Education, Cooperative Research Project No. 2410. Durant, Oklahoma, 1965.

Strickland, Ruth G. The Language of Elementary School Children: Its Relationship to the Language of Reading Textbooks and the Quality of Reading of Selected Children. Bulletin of the School of Education, Indiana University, 38, No. 4. Bloomington: Indiana University, 1962.

Veal, L. Ramon. "Syntactic Measures and Rated Quality in the Writing of Young Children." Unpublished paper, available from the author, University of Georgia, 1969.